

Technická univerzita v Liberci

**FAKULTA PŘÍRODOVĚDNĚ-HUMANITNÍ
A PEDAGOGICKÁ**

Katedra: Českého jazyka a literatury

Studijní program: Specializace v pedagogice

Studijní obor: Český jazyk a literatura - Anglický jazyk a literatura

Bakalářská práce:

**Porovnání hranic větných celků v závislosti na
realizační formě textu**

**Comparison of the Boundaries of Sentence Units,
Depending on the Form of Realisation**

Autor:

Hana Plocová

Podpis:

Vedoucí práce:

Mgr. Svatava Škodová, PhD.

Počet

stran	grafů	obrázků	tabulek	pramenů	příloh
59	0	2	29	12	2

V Liberci dne:

Čestné prohlášení

Název práce:	Porovnání hranic větných celků v závislosti na realizační formě textu
Jméno a příjmení autora:	Hana Plocová
Osobní číslo:	P08000279

Byla jsem seznámena s tím, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, zejména § 60 – školní dílo.

Prohlašuji, že má bakalářská práce je ve smyslu autorského zákona výhradně mým autorským dílem.

Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé bakalářské práce pro vnitřní potřebu TUL.

Užiji-li bakalářskou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědoma povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Bakalářskou práci jsem vypracovala samostatně s použitím uvedené literatury a na základě konzultací s vedoucím bakalářské práce a konzultantem.

Prohlašuji, že jsem do informačního systému STAG vložila elektronickou verzi mé bakalářské práce, která je identická s tištěnou verzí předkládanou k obhajobě a uvedla jsem všechny systémem požadované informace pravdivě.

V Liberci dne:

Hana Plocová

Anotace

Bakalářská práce zkoumá umístění hranic větných celků v závislosti na realizační formě textu, kterou je buď mluvený, nebo psaný projev. Skupina respondentů má za úkol postupně do obou forem doplnit interpunkční znaménka. Vyhodnocuje se, která znaménka a na jakém místě v textu byla použita. Důraz je kladen na porovnání výsledků v rámci jednoho respondenta i srovnání respondentů mezi sebou.

Bakalářská práce využívá dvou nástrojů vyvinutých na Technické univerzitě v Liberci, Nanotransu – rozpoznávače řeči, a Transcript Vieweru – porovnávače rozpoznávaných textů.

Klíčová slova: interpunkce, znaménka, větné celky, Nanotrans, Transcript Viewer

The bachelor thesis examines the position of sentence units' borders, depending on the form of text realisation. The form is either spoken or written. The group of respondents is instructed to fill in punctuation marks to both forms. It is evaluated which punctuation marks and on which position has been used. The emphasis is given to the comparison of results of one respondent as well as to comparison of respondents with each other.

The bachelor thesis uses two tools developed at Technical university of Liberec. Nanotrans – the recognition tool, and Transcript Viewer – the comparator of recognised texts.

Key words: punctuation mark, sentence units, Nanotrans, Transcript Viwer

Obsah

1	Seznam obrázků.....	5
2	Seznam tabulek.....	5
3	Úvod	6
4	Interpunkce	8
5	Přípravné práce	9
5.1	Nanotrans	9
5.1.1	Příprava materiálu v Nanotransu	11
5.2	Výběr respondentů	11
5.3	Zadání práce respondentům.....	11
6	Shromažďování dat.....	13
7	Práce s daty	14
7.1	Transcript Viewer (Porovnávač textů)	14
8	Vyhodnocení dat.....	16
8.1	Doplnění interpunkce	16
8.1.1	Součty	17
8.1.2	Tečky	20
8.1.3	Čárky	26
8.1.4	Otazníky.....	32
8.1.5	Vykřičníky	38
8.2	Porovnání s automatikou	40
8.2.1	Automatika obecně	40
8.2.2	Automatika konkrétně	42
9	Závěr	53
10	Seznam literatury	54
11	Přílohy	55

1 Seznam obrázků

Obrázek 1: Nanotrans 10

Obrázek 2: Transcript Viewer (Porovnávač textů)..... 14

2 Seznam tabulek

Tabulka 1: Shoda mluvčích ve všech interpunkčních znaménkách 16

Tabulka 2: Úsek 1 součty 17

Tabulka 3: Úsek 2 součty 17

Tabulka 4: Úsek 3 součty 18

Tabulka 5: Úsek 4 součty 18

Tabulka 6: Úsek 5 součty 18

Tabulka 7: Úsek 1 tečky 21

Tabulka 8: Úsek 2 tečky 22

Tabulka 9: Úsek 3 tečky 23

Tabulka 10: Úsek 4 tečky 24

Tabulka 11: Úsek 5 tečky 25

Tabulka 12: Úsek 1 čárky 27

Tabulka 13: Úsek 2 čárky 28

Tabulka 14: Úsek 3 čárky 29

Tabulka 15: Úsek 4 čárky 30

Tabulka 16: Úsek 5 čárky 31

Tabulka 17: Úsek 1 otazníky 33

Tabulka 18: Úsek 2 otazníky 34

Tabulka 19: Úsek 3 otazníky 35

Tabulka 20: Úsek 4 otazníky 36

Tabulka 21: Úsek 5 otazníky 37

Tabulka 22: Úsek 1-5 vykřičníky 39

Tabulka 23: Automatika 40

Tabulka 24: Respondenti 41

Tabulka 25: Úsek 1: Porovnání automatiky a respondentů 43

Tabulka 26: Úsek 2: Porovnání automatiky a respondentů 45

Tabulka 27: Úsek 3: Porovnání automatiky a respondentů 47

Tabulka 28: Úsek 4: Porovnání automatiky a respondentů 49

Tabulka 29: Úsek 5: Porovnání automatiky a respondentů 51

3 Úvod

Předmětem zkoumání této bakalářské práce je umístění hranic větných celků v závislosti na realizační formě textu. Text tvoří pomocí počítače rozpoznané a přepsané rozhlasové pořady Leonardo, z nichž byl vybrán reprezentativní vzorek mluvčích tak, aby bylo možné případně porovnávat schopnost recipientů rozpoznat hranice vět u profesionálních a neprofesionálních mluvčích. Realizační formou textu v našem případě rozumíme jednak čistě textovou podobu a jednak text se zvukovou oporou.

Zjišťujeme, zda realizační forma zásadně ovlivní vnímání recipientů, jejichž úkolem je doplnit interpunkci do textu, a zda se budou respondenti lišit nejen mezi sebou, ale i jedna osoba v rámci svých odpovědí bude vykazovat odchylky v pochopení jednotlivých hranic větných celků u dvou typů realizačních forem.

Předpokládáme, že shoda v doplněné interpunkci bude výrazně větší u textu se zvukovou oporou, protože mluvčí svým projevem, intonací a pauzami poskytne recipientům důležitá vodítka pro rozhodnutí, kam interpunkci doplnit.

V textu *K základním otázkám interpunkce v češtině* od Miloslava Sedláčka se o vztahu interpunkce a intonace dočteme následující: „*Interpunkce má v psaném (tištěném) textu zčásti obdobnou funkci jako intonační kadence a pauzy v projevu mluveném, avšak důsledný funkční paralelismus mezi interpunkcí a zvukovým členěním není.*“¹ V duchu tohoto tvrzení ověřujeme, zda intonační ztvárnění (pronesení)² se nějakým způsobem promítne i do přepisu mluveného slova a respondenti rozeznají původní hranice větných celků.

Jedna z kapitol bude věnována také porovnání počítačem automaticky doplněné interpunkce s texty dodanými respondenty. Zkoumat se bude množství shod v užití znamének jak do typu, tak umístění.

Bakalářská práce je součástí projektu *Zpřístupnění archivu Českého rozhlasu pro sofistikované vyhledávání* (DF11P010VV013, podpořeného Ministerstvem kultury ČR ve výzkumném programu NAKI). Získaná data by měla posloužit jako opora při

¹ SEDLÁČEK, Miloslav. K základním otázkám interpunkce v češtině. *Naše řeč*. 1986, roč. 69, č. 3, s. 121-132. ISSN 0027-8203. S. 121.

² GREPL, Miroslav a kol.: Příruční mluvnice češtiny. Praha : NLN, 2008. ISBN 978-80-7106-624-8

kvantitativní, počet interpunkčních znamének, a kvalitativní, druh interpunkčních znamének, analýze pomocí počítače automaticky doplňovaných jevů.

Posouzení shody mezi mluvčími dodá jasnější představu o odlišnosti chápání hranic větných celků při zvukové opoře a bez ní. Srovnání automatického doplnění interpunkce s výsledky respondentů zdůrazní oblasti, kde jsou zatím možnosti počítače omezeny, a oblasti, kde preciznost doplnění interpunkčních znamének, kterou disponuje počítač (v místech jasně gramaticky vymezených, např. před spojkami), předčí lidský faktor, náchylný k chybovosti z nedbalosti.

4 Interpunkce

Definici slova interpunkce nám poskytne Miloslav Sedláček: „*Interpunkcí se rozumí členění psaného (tištěného) textu a jeho složek (výpovědí) podle ustálených (mnohdy kodifikovaných) zásad a pomocí souboru zvláštních grafických znamének.*“³

Oblast interpunkce je velmi důležitou součástí všech psaných forem jazyka díky své schopnosti text členit a dávat mu jasnou strukturu. Mimo tuto funkci ovlivňuje doplnění či nedoplnění interpunkce i samotný význam vět.

Znaménka na koncích vět jasně odkazují ke komunikační funkci výpovědi, k tomu, zda se jedná o větu oznamovací, zpravidla ukončenou tečkou, rozkazovací, ukončenou vykřičníkem, nebo tázací, na konci s otazníkem. Mezi základní čtyři typy vět podle komunikační funkce řadíme ještě věty přací, zakončené tečkou nebo vykřičníkem.

Čárky a jejich sémantická funkce bývá často podceňována. Např. vložené vedlejší věty, v jejichž případě většinou, ne ovšem vždy, při doplňování interpunkce vystupuje do popředí převážně požadavek na strukturaci textu, bývají často spílaným problémem. Správné očárkování se nezdá být dostatečně důležité. V mnoha případech ale čárky zásadně ovlivňují význam vět a jejich doplnění či nedoplnění není pouze otázkou přehlednosti textu. Typickým a s oblibou uváděným příkladem je věta „Propustit (,) nelze (,) popravit.“, kde čárka skutečně a doslova rozhoduje o životě a smrti.

Ne všechny případy (ne)doplnění čárek jsou však fatální. Přesto je zřejmé, že musí existovat místa v textu, kdy ani rodilí mluvčí, zběhlí v lingvistice, nejsou schopni bez dalších souvislostí čárky doplnit „správně“. Gramatická ani významová správnost není vždy jednoznačná a varianty doplnění si mohou být zcela rovnocenné.

³ SEDLÁČEK, Miloslav. K základním otázkám interpunkce v češtině. *Naše řeč*. 1986, roč. 69, č. 3, s. 121-132. ISSN 0027-8203. S. 121.

5 Přípravné práce

Před započítím samotného výzkumu bylo třeba vytvořit jasnou strukturu a vymyslet způsob, jak data od respondentů dostat v požadované podobě a relevantní formě.

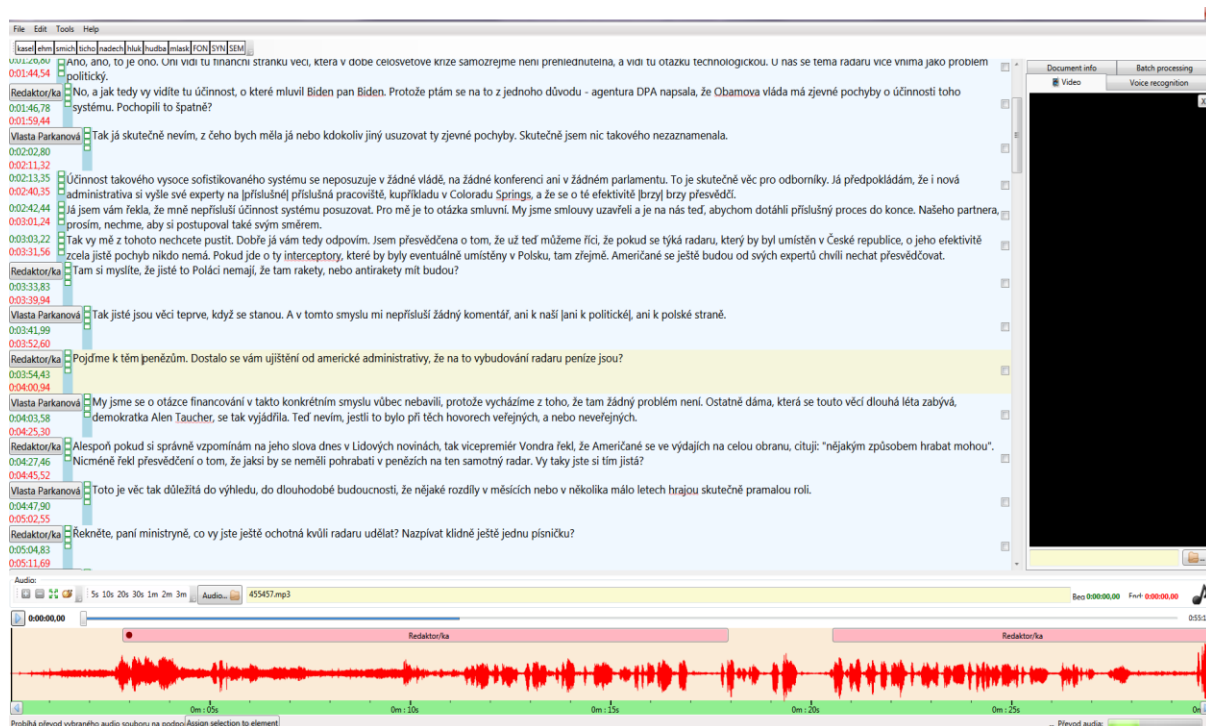
Přípravné práce probíhaly s využitím různých nástrojů, speciálně určených k práci s texty tvořenými počítačově rozpoznávaným zvukovým materiálem.

5.1 Nanotrans

Nanotrans je program vyvinutý v Ústavu informačních technologií a elektroniky (ITE) na Technické univerzitě v Liberci. Slouží k počítačovému přepisu mluveného slova do textové podoby s možností vyznačení mimojazykových jevů (smích, kašel, hesitační zvuky). Pomocí schopnosti rozlišovat jednotlivé mluvčí rozdělí Nanotrans text do kratších segmentů, kdy každý promluvový akt jednoho mluvčího odpovídá jednomu segmentu.

Lexikální správnost rozpoznávaného textu je velmi vysoká, i když se zatím neobejde bez následné korektury. Je zřejmé, že procentuální úspěšnost správně přepsaného textu závisí především na konkrétním mluvčím a jeho dovednosti artikulace. Přepisy mluvených projevů některých mluvčích dosahují velmi vysoké úspěšnosti, jiné, od méně dovedných rétorů, naopak dělají Nanotransu problém.

V zatím omezené míře je program schopen doplnit do vzniklého textu interpunkci. Čárky jsou bezchybně doplňovány hlavně před podřadné spojky (který, že, protože ...), problémy se objevují v koncích vedlejších vět, kde čárky převážně chybí, a u oslovení, jež zatím Nanotrans není schopen rozeznat.



Obrázek 1: Nanotrans

5.1.1 Příprava materiálu v Nanotransu

Automaticky rozpoznáný text bylo třeba kvůli výše uvedeným problémům opravit. Nejprve se pomocí zvukové stopy kontrolovala lexikální správnost, morfologická přesnost a nakonec byl důraz kladen i na gramatiku. Důležité bylo především uvést do pořádku shodu podmětu s přísudkem, aby tyto chyby v dalších fázích respondenty nemátly, a doplnit interpunkci na patřičná místa.

Interpunkce byla doplněna pouze pro snazší orientaci korektora v textu, neboť před předáním dat respondentům byla vymazána nejen všechna znaménka, ale i veškerá velká písmena, aby nedošlo k nechtěnému ovlivnění respondentů.

5.2 Výběr respondentů

Skupinu respondentů tvořili studenti a studentky katedry Českého jazyka a literatury (KČL) na Fakultě přírodovědně-humanitní a pedagogické (PF) Technické univerzity v Liberci (TUL). Záměrně byli do výzkumu zařazeni bohemisté, protože u nich se předpokládá znalost českého jazyka a jeho zákonitostí na velmi vysoké úrovni, a je proto pravděpodobné, že interpunkce se bude většinou lišit pouze na místech nejasností a lépe vyniknou varianty hranic vět na specifických místech textu.

Původní počet respondentů byl stanoven na padesát. Vlivem vnějších okolností, nesprávného odevzdání zpracovaného úkolu a jiných komplikací se nakonec počet, se kterým se dále pracovalo, snížil na 45.

5.3 Zadání práce respondentům

Všichni respondenti obdrželi balíček s programem Nanotrans, v němž bylo nutné úkol zpracovávat, dále Profil respondenta a Průvodce jednotlivými kroky.

Průvodce posloužil jako základní pomůcka při práci s Nanotransem, protože žádný ze studentů se předtím s tímto programem nesetkal, a jako zadavatel konkrétních úkolů.

Úkolem respondentů bylo do textu, z něhož byla záměrně odstraněna veškerá interpunkce, doplnit gramaticky a sémanticky správně hranice větných celků. K dispozici dostali všechna znaménka, objektem zájmu této bakalářské práce ovšem byla pouze následující: otazník (?), tečka (.), čárka (,) a vykřičník (!). Užívání jiných interpunkčních znamének ani označování přímé řeči nebylo součástí výzkumu.

Nejprve byla znaménka doplněna do textu bez toho, aby respondenti měli možnost porovnat přepsaný text se zvukovým záznamem. Po odevzdání této části jim Průvodce doporučoval minimálně třídní rozestup před zpracováváním další části. Časovým rozestupem se snížilo riziko ovlivnění předchozím doplněním interpunkce.

Druhá část byla zadáním velmi obdobná první části, tentokrát ovšem měli respondenti za úkol spustit k doplňování i zvukovou stopu. Jejich rozhodování tak bylo, alespoň dle předpokladů, ovlivněno dikcí a intonací mluvčího.

6 Shromažďování dat

Data byla respondenty odevzdávána elektronickou formou a po základním posouzení jejich validity byla zařazena do výzkumu. Vyřazena musela být například data se špatně vepsanou interpunkcí nebo ta, která neměla jednu z požadovaných forem, tj. byl odevzdán buď pouze zvukový přepis, nebo textový bez zvukové opory.

Dále zpracovávána tedy nakonec byla data od 45 respondentů – 45krát text, do něhož byla interpunkce doplněna bez zvukové opory, 45krát text s interpunkcí doplněnou pomocí zvukové stopy a 45krát Profil respondenta, krátký anonymní dotazník, který poskytl základní informace o osobě, jež úkol zpracovávala. Profil respondenta nebyl pro účely bakalářské práce použit, dává tedy prostor pro případná další zkoumání dané problematiky.

7 Práce s daty

Získaná data bylo nutné zpracovávat v programu, který by byl schopen rozdělit text na takové segmenty, aby se dalo porovnávat umístění interpunkce na konkrétních místech v textu. K tomuto účelu byl vyvinut speciální program, který pracuje s daty získanými v Nanotransu, Transcript Viewer.

7.1 Transcript Viewer (Porovnávač textů)

Autorkou Transcript Vieweru je Ing. Michaela Kuchařová z katedry Informačních technologií, s jejíž pomocí byl program v průběhu zpracovávání dat konzultován a k účelům bakalářské práce doplněn o některé funkce.

Na Obrázku 1 vidíme rozložení dat v Transcript Vieweru. V levém horním rohu nalezneme informace o aktuální stránce, se kterou pracujeme, její pořadí a jméno mluvčího. Ve sloupečku po levé straně jsou seřazeni jednotliví respondenti a v příslušných řádcích je jimi vyplněný text. Barevně program vyznačuje doplněnou interpunkci.

Transcript Viewer																																		
Actual position is: 25 from 81 blocks. Speaker is: Pavel Telicka																																		
Text 32 10/80	do	ještě	mily	si	získal	jméno	,	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 33 9/79	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 34 10/80	do	ještě	mily	si	získal	jméno	,	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 35 11/81	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 36 11/81	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 37 13/83	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 38 13/83	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 39 10/80	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 40 12/82	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 41 14/84	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 42 13/83	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 43 11/81	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 44 10/80	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Text 45 14/84	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Zvuk 01 0/70	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Zvuk 02 8/78	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Zvuk 03 12/82	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Zvuk 04 10/80	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Zvuk 05 11/81	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Zvuk 06 10/80	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Zvuk 07 11/81	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale
Zvuk 08 11/81	do	ještě	mily	si	získal	jméno	.	alespoň	je	znám	,	a	on	rovněž	zná	případně	potenciální	patrně	.	na	druhou	stranu	samožřejmě	je	to	člověk	,	kteř	má	i	určitý	handicap	,	ale

Obrázek 2: Transcript Viewer (Porovnávač textů)

Transcript Viewer není zatím schopný data mezi sebou aktivně porovnávat, řadit je podle určitého klíče ani nedisponuje pro potřeby této bakalářské práce velmi důležitým sčítáním jednotlivých jevů. Data, Transcript Viewerem seřazená do požadované podoby, byla tedy postupně zkopírována do programu Microsoft Excel, který nabízí dostatečné funkce pro pokročilou práci s datovým materiálem.

Pomocí funkcí, které vyfiltrovaly všechna slova, byla v textovém materiálu vyhledávána interpunkční znaménka. Objem dat neumožňoval zkoumat celý text jako celek, proto byl rozdělen na 5 částí, Úseků. Klíčovým kritériem při dělení do Úseků byli mluvčí, jeden Úsek tedy zaznamenává rozhlasový rozhovor s jediným mluvčím.

Pro každý Úsek vznikla tabulka, v níž se užitím dalších funkcí zobrazovaly počty jednotlivých znamének (tečka, čárka, otazník, vykřičník) v celé délce Úseku.

Postup se uplatňoval u textové i zvukové podoby. Následovalo srovnání, zda na přesném místě v textu i zvuku doplnil respondent stejné znaménko.

8 Vyhodnocení dat

Při vyhodnocování dat musíme mít neustále na paměti fakta, o kterých se píše v kapitole č. 2 této bakalářské práce. Existují místa v textu, kde jednoznačné doplnění interpunkce není z určitých důvodů možné. Ne všechny rozdíly v doplnění interpunkce nutně znamenají, že některý z respondentů se mýlí a jiný má pravdu. Naopak, místa s výraznými odlišnostmi u jednotlivých respondentů jsou statisticky nejzajímavější, protože pomáhají pochopit variantnost češtiny a shrnout všechny odlišnosti a možnosti pochopení jedné věty.

8.1 Doplnění interpunkce

Při hodnocení dat pozornost směřovala především k porovnání interpunkce, doplněné bez zvukové stopy (dále textová podoba, Text) a následně se zvukovou stopou (dále zvuková podoba, Zvuk), u jednoho respondenta.

Následující tabulka je shrnutím celkové sumy obdržených informací.

Tabulka 1: Shoda mluvčích ve všech interpunkčních znaménkách

Číslo úseku	Shoda mluvčích
Úsek 1	62%
Úsek 2	42%
Úsek 3	22%
Úsek 4	53%
Úsek 5	50%

Úsek 1 až 5 označuje celky, ze kterých se přepisovaný text skládá. *Úsek 1* je rozhovorem Redaktorky 1 s Vlastou Parkanovou, *Úsek 2* Redaktor 1 a Pavel Telička, *Úsek 3* Redaktor 2 a Daniela Filipi, *Úsek 4* Redaktorka 2 a Radek John a *Úsek 5* Redaktor 3 a Zdeněk Škromach.

Procenta udávají shodu v umístění interpunkčního znaménka na jednom místě v textu. Ta se počítá z procentuální úspěšnosti jednoho respondenta udat stejné znaménko do své textové a zvukové podoby textu, načež je procentuální shoda u jednotlivých znamének zprůměrována. Výsledná hodnota je pak použita při počítání konečného výsledku v rámci ostatních respondentů.

Zprůměrováním v tabulce uvedených údajů dojdeme k celkové shodě 46 %, což značí shodu v méně než polovině znamének.

Na následujících stránkách se budeme podrobně zabývat jednotlivými znaménky, zkoumat shodu v jejich doplnění a celkový počet doplněných znamének daného druhu. Předpokládáme, že větší množství znamének bude respondenty použito u textové podoby. K tomuto tvrzení dojdeme, uvědomíme-li si, že při čtení je nejprve nutné správně se v neznámém textu zorientovat. Potřeba segmentace a strukturace je tedy v danou chvíli vyšší a respondenti se snaží text rozdělit na kratší a srozumitelnější úseky.

8.1.1 Součty

Všechna sebraná data jsou přehledně seřazena do pěti tabulek. Nahoře vidíme číslo úseku, kterého se tabulka týká, a v jednotlivých sloupcích nacházíme postupně druh znaménka (*Znaménko*), počet doplnění příslušných znamének v pracích všech respondentů pro textovou (Text) a zvukovou podobu (Zvuk), shodu v umístění daného znaménka (*Počet shod Text + Zvuk*) a shodu vyjádřenou procentuální hodnotou (*% shody*).

Tabulka 2: Úsek 1 součty

Úsek 1				
Znaménko	Text	Zvuk	Počet shod Text + Zvuk	% shody
.	1432	1316	1169	85%
,	2797	2766	2250	81%
!	10	5	2	27%
?	419	407	337	82%
Celkem	4658	4494	3758	69%

Tabulka 3: Úsek 2 součty

Úsek 2				
Znaménko	Text	Zvuk	Počet shod Text + Zvuk	% shody
.	1313	1295	802	62%
,	3781	3770	2562	68%
!	2	1	1	67%
?	167	158	58	36%
Celkem	5263	5224	3423	58%

Tabulka 4: Úsek 3 součty

Úsek 3				
Znaménko	Text	Zvuk	Počet shod Text + Zvuk	% shody
.	1031	773	438	49%
,	3742	3231	1421	41%
!	4	1	0	0%
?	244	205	1	0%
Celkem	5021	4210	1860	22%

Tabulka 5: Úsek 4 součty

Úsek 4				
Znaménko	Text	Zvuk	Počet shod Text + Zvuk	% shody
.	1788	1329	1013	65%
,	4345	3390	2873	74%
!	27	15	11	52%
?	382	278	192	58%
Celkem	6542	5012	4089	62%

Tabulka 6: Úsek 5 součty

Úsek 5				
Znaménko	Text	Zvuk	Počet shod Text + Zvuk	% shody
.	1604	1092	848	63%
,	4416	3241	2654	69%
!	10	7	3	35%
?	225	133	117	65%
Celkem	6255	4473	3622	58%

Jak vidíme, potvrdil se výše zmíněný předpoklad, že v textové podobě respondenti užijí více interpunkčních znamének než v podobě zvukové. V Úseku 2 se jedná o rozdíl 39 znamének, Úsek č. 5 se liší dokonce v 1782 znaménkách.

Každý Úsek je ovšem specifický v tom, které znaménko ve Zvuku snižovalo svůj počet. V Úseku č. 1 je víceméně zachován počet čárek, vykřičníků a otazníků, nejvýrazněji se mění počet teček, z původních 1432 na 1316. Oproti tomu v Úseku č. 2 se změna týkala teček i čárek v přibližně stejném měřítku (18 teček a 11 čárek).

Úsek 3 vykazuje rozdíl v 258 tečkách a 511 čárkách. Úsek 4 se liší v 459 tečkách a 955 čárkách. Úsek 5 v 512 tečkách a 1175 čárkách.

Výjimečnost Úseku 2 spočívající ve schopnosti respondentů shodnout se na doplněné interpunkci v Textu i Zvuku ukazuje předně na dovednost mluvčích přizpůsobit svůj mluvený projev psané, kodifikaci dané normě. Respondenti byli schopni rozeznat hranice větných celků i v do textu převedené podobě mluveného slova. Dalším důvodem pro nižší početní rozdíl mohou být také zvolené jazykové prostředky. Méně časté užívání rozvinutých a několikanásobných přívlastků nebo snížený výskyt podřadných spojek, které by vytvářely delší, zřetězené větné celky. Na takových místech záleží čistě na úvaze respondenta, kdy by už bylo vhodné větu ukončit.

Úseky 1, 3, 4 a 5, v nichž se pravděpodobně objevili méně zkušené rétoři, ukazují na problémy s porozuměním textu. Jeho nejednoznačnost vedla respondenty k tomu, že se snažili v text bez zvukové opory co nejvýrazněji rozčlenit. Na větě z Úseku 5 si můžeme doložit varianty možných doplnění.

Věta bez zvukového záznamu vypadá takto: „*určitě je to sféra bankovníctví kde se řekněme ne zrovna zodpovědně nakládalo s financemi úvěrama*“, přičemž víme, že věta je odpovědí na předchozí otázku, vyloučíme tedy doplnění otazníků.

Zůstávají nám tedy tyto, gramaticky správné, možnosti:

- 1) *Určitě! Je to sféra bankovníctví, kde se řekněme ne zrovna zodpovědně nakládalo s financemi, úvěrama.*
- 2) *Určitě. Je to sféra bankovníctví, kde se řekněme ne zrovna zodpovědně nakládalo s financemi, úvěrama.*
- 3) *Určitě je to sféra bankovníctví, kde se řekněme ne zrovna zodpovědně nakládalo s financemi, úvěrama.*

Za slovem *určitě* může ve větě následovat buď vykřičník (1), tečka (2), nebo, chápeme-li slovo *určitě* jako příslovečné určení, prázdné okénko (3).

K dalšímu problému bychom se dostali při úvaze, jestli oddělit, nebo neoddělit z obou stran čárkami také slovo *řekněme*. Většina respondentů se rozhodla pro oddělení v Textu a neoddělení ve Zvuku, což jednoznačně ukazuje, že hlasový projev mluvčího zásadně ovlivnil jejich rozhodnutí.

8.1.2 Tečky

Níže uvedené tabulky nás seznamují s výsledky získanými pro oblast teček. V prvním sloupečku vidíme čísla, která byla přiřazena jednotlivým respondentům, ve druhém a třetím počet teček doplněný daným respondentem do jeho textové (*Text*) a zvukové podoby (*Zvuk*) materiálu. Třetí sloupec ukazuje shodu v druhu znaménka a přesném umístění v textu u *Textu* a *Zvuku* u jednoho respondenta.

Například v *Úseku 1* doplnil *Respondent 8* do *Textu* 42 teček a do *Zvuku* 31 teček. Co do přesného umístění (*Počet shod*) se shodovalo 27 čárek, což činí celkovou shodu (*% shody*) 74 %.

V celkovém průměru můžeme říct, že teček bylo doplněno více do *Textu* než do *Zvuku* (pravděpodobné důvody jsou popsány v předchozí kapitole), u jednotlivých respondentů se ale setkáváme i se situací, kdy je více teček doplněno do *Zvuku*.

Tabulka 7: Úsek 1 tečky

Respondent	Počet teček	Počet teček	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	35	29	26	81%
2	32	37	30	87%
3	38	31	27	78%
4	29	25	23	85%
5	33	32	30	92%
6	28	29	25	88%
7	28	33	27	89%
8	32	32	28	88%
9	30	33	27	86%
10	33	28	26	85%
11	25	30	22	80%
12	32	30	27	87%
13	21	20	14	68%
14	36	32	29	85%
15	30	35	22	68%
16	29	25	23	85%
17	36	25	24	79%
18	25	32	22	77%
19	32	33	29	89%
20	34	34	30	88%
21	24	20	18	82%
22	34	28	24	77%
23	38	31	29	84%
24	32	17	16	65%
25	31	30	29	95%
26	37	34	32	90%
27	20	18	13	68%
28	30	14	14	64%
29	36	35	35	99%
30	34	36	32	91%
31	24	24	24	100%
32	35	34	34	99%
33	39	38	35	91%
34	34	32	32	97%
35	33	22	22	80%
36	25	20	17	76%
37	31	32	27	86%
38	32	23	21	76%
39	31	33	31	97%
40	37	36	34	93%
41	36	25	24	79%
42	33	36	28	81%
43	37	36	36	99%
44	32	25	21	74%
45	39	32	30	85%

Tabulka 8: Úsek 2 tečky

Respondent	Počet teček	Počet teček	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	29	22	11	43%
2	28	41	17	49%
3	34	46	24	60%
4	27	29	14	50%
5	29	28	19	67%
6	22	26	16	67%
7	24	24	11	46%
8	27	36	22	70%
9	23	40	19	60%
10	27	23	16	64%
11	26	36	20	65%
12	24	25	16	65%
13	21	18	10	51%
14	31	24	15	55%
15	38	42	26	65%
16	24	21	15	67%
17	38	32	20	57%
18	21	26	14	60%
19	32	36	17	50%
20	33	29	17	55%
21	28	12	10	50%
22	25	26	17	67%
23	26	38	17	53%
24	23	10	7	42%
25	25	26	16	63%
26	30	35	21	65%
27	29	19	15	63%
28	30	16	11	48%
29	38	38	30	79%
30	31	37	23	68%
31	21	23	13	59%
32	28	28	23	82%
33	41	39	26	65%
34	28	28	23	82%
35	29	20	13	53%
36	30	19	14	57%
37	32	31	20	63%
38	28	31	15	51%
39	28	31	24	81%
40	36	38	26	70%
41	46	25	17	48%
42	25	29	15	56%
43	38	38	33	87%
44	23	18	8	39%
45	37	36	26	71%

Tabulka 9: Úsek 3 tečky

Respondent	Počet teček	Počet teček	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	23	13	10	56%
2	24	29	12	45%
3	36	35	18	51%
4	20	18	9	47%
5	19	14	7	42%
6	21	17	7	37%
7	20	12	8	50%
8	22	19	11	54%
9	21	24	10	44%
10	22	12	8	47%
11	14	17	8	52%
12	20	11	7	45%
13	18	12	9	60%
14	23	13	7	39%
15	30	24	9	33%
16	26	12	9	47%
17	28	20	11	46%
18	19	18	9	49%
19	19	21	8	40%
20	24	14	9	47%
21	14	11	6	48%
22	20	15	9	51%
23	21	18	9	46%
24	15	8	7	61%
25	21	15	11	61%
26	30	23	13	49%
27	20	12	7	44%
28	24	9	7	42%
29	28	22	15	60%
30	30	27	15	53%
31	20	15	11	63%
32	17	11	8	57%
33	34	24	14	48%
34	17	11	8	57%
35	26	19	15	67%
36	15	14	8	55%
37	25	15	8	40%
38	25	25	11	44%
39	28	16	11	50%
40	28	24	14	54%
41	38	13	8	31%
42	22	16	7	37%
43	23	17	10	50%
44	14	13	7	52%
45	27	25	13	50%

Tabulka 10: Úsek 4 tečky

Respondent	Počet teček	Počet teček	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	32	26	19	66%
2	40	43	27	65%
3	50	46	34	71%
4	29	33	17	55%
5	41	30	27	76%
6	35	23	14	48%
7	40	27	20	60%
8	41	28	22	64%
9	41	37	24	62%
10	42	30	23	64%
11	35	33	20	59%
12	31	27	16	55%
13	24	21	13	58%
14	41	32	20	55%
15	58	44	29	57%
16	40	26	21	64%
17	52	34	27	63%
18	36	30	21	64%
19	45	33	23	59%
20	40	22	19	61%
21	23	24	13	55%
22	43	28	20	56%
23	38	32	20	57%
24	28	12	9	45%
25	35	28	25	79%
26	41	37	28	72%
27	34	18	13	50%
28	34	18	16	62%
29	44	33	33	86%
30	48	38	29	67%
31	41	25	24	73%
32	30	23	23	87%
33	59	37	34	71%
34	32	22	22	81%
35	45	33	32	82%
36	29	20	15	61%
37	46	32	26	67%
38	39	27	19	58%
39	39	29	25	74%
40	48	35	29	70%
41	53	26	21	53%
42	44	31	26	69%
43	49	34	34	82%
44	31	26	16	56%
45	42	36	25	64%

Tabulka 11: Úsek 5 tečky

Respondent	Počet teček	Počet teček	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	30	29	11	37%
2	42	36	27	69%
3	42	41	26	63%
4	27	28	15	55%
5	34	18	16	62%
6	32	26	15	52%
7	31	21	18	69%
8	42	31	27	74%
9	32	26	20	69%
10	33	22	18	65%
11	38	34	24	67%
12	29	19	16	67%
13	23	15	12	63%
14	37	20	14	49%
15	50	40	26	58%
16	36	23	18	61%
17	57	30	27	62%
18	28	26	18	67%
19	36	24	20	67%
20	38	16	16	59%
21	24	17	12	59%
22	41	20	17	56%
23	32	25	17	60%
24	20	11	10	65%
25	37	25	23	74%
26	43	30	24	66%
27	26	16	13	62%
28	32	12	10	45%
29	42	30	27	75%
30	35	31	23	70%
31	33	20	20	75%
32	32	34	18	55%
33	53	32	25	59%
34	28	20	18	75%
35	6	15	6	57%
36	24	18	12	57%
37	45	27	24	67%
38	47	17	14	44%
39	42	26	25	74%
40	42	29	27	76%
41	51	20	18	51%
42	42	22	19	59%
43	41	28	27	78%
44	24	15	13	67%
45	45	27	22	61%

8.1.3 Čárky

Čárky se v rámci výzkumu ukázaly být nejpočetněji doplňovaným interpunkčním znaménkem. Tabulky na následujících stránkách nám předkládají získané údaje. Rozvržení tabulek je stejné jako v předchozí podkapitole, v prvním sloupečku vidíme číslo respondenta, ve druhém a třetím doplnění čárek pro jednotlivé druhy materiálu, ve čtvrtém počet shod u jednotlivce pro Text a Zvuk a v posledním sloupci procentuálně vyjádřenou shodu.

Například v *Úseku 2* doplnil *Respondent 24* do *Textu* 84 čárek, do *Zvuku* 99 čárek, z čehož se v 61 případech shodl v jejich umístění (*Počet shod*). Celkem tedy shoda činila 67 % (*% shody*).

Tabulka 12: Úsek 1 čárky

Respondent	Počet teček	Počet teček	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	92	94	73	78%
2	43	43	30	70%
3	61	46	38	71%
4	58	52	46	84%
5	56	59	50	87%
6	59	56	44	77%
7	52	60	38	68%
8	65	61	57	90%
9	56	62	48	81%
10	67	69	58	85%
11	64	54	49	83%
12	62	66	51	80%
13	63	60	47	76%
14	57	63	50	83%
15	66	53	42	71%
16	66	68	56	84%
17	61	67	42	66%
18	70	58	53	83%
19	58	54	49	88%
20	60	60	48	80%
21	55	68	50	81%
22	63	56	42	71%
23	67	63	51	78%
24	59	79	54	78%
25	65	68	63	95%
26	69	66	60	89%
27	66	64	45	69%
28	65	69	46	69%
29	62	62	61	98%
30	57	53	49	89%
31	63	69	61	92%
32	71	69	67	96%
33	41	36	32	83%
34	69	74	67	94%
35	63	56	42	71%
36	69	70	59	85%
37	71	68	58	83%
38	60	51	26	47%
39	57	63	56	93%
40	65	49	47	82%
41	53	69	47	77%
42	68	61	45	70%
43	57	57	56	98%
44	57	57	40	70%
45	69	64	57	86%

Tabulka 13: Úsek 2 čárky

Respondent	Počet čárek	Počet čárek	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	92	71	44	54%
2	64	59	38	62%
3	88	76	59	72%
4	85	75	55	69%
5	81	86	60	72%
6	80	84	58	71%
7	69	93	49	60%
8	89	85	68	78%
9	88	86	57	66%
10	85	91	61	69%
11	78	69	47	64%
12	88	89	63	71%
13	82	69	49	65%
14	82	94	62	70%
15	77	69	42	58%
16	83	84	57	68%
17	98	101	65	65%
18	95	89	69	75%
19	82	71	49	64%
20	83	83	54	65%
21	77	80	54	69%
22	83	91	56	64%
23	100	85	61	66%
24	84	99	61	67%
25	90	89	74	83%
26	90	82	62	72%
27	82	76	47	59%
28	83	98	56	62%
29	91	91	71	78%
30	74	77	54	72%
31	91	99	72	76%
32	90	90	70	78%
33	56	55	39	70%
34	90	90	70	78%
35	88	73	39	48%
36	81	98	62	69%
37	94	89	65	71%
38	83	68	28	37%
39	81	78	60	75%
40	80	84	52	63%
41	78	100	57	64%
42	89	96	63	68%
43	78	78	62	79%
44	90	92	54	59%
45	89	88	67	76%

Tabulka 14: Úsek 3 čárky

Respondent	Počet čárek	Počet čárek	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	87	68	34	44%
2	65	49	23	40%
3	82	68	27	36%
4	62	67	23	36%
5	79	75	33	43%
6	82	60	29	41%
7	73	73	27	37%
8	84	71	36	46%
9	94	87	32	35%
10	88	80	34	40%
11	77	57	26	39%
12	89	74	33	40%
13	72	58	27	42%
14	86	68	27	35%
15	71	56	16	25%
16	78	66	32	44%
17	95	84	36	40%
18	85	84	32	38%
19	90	72	33	41%
20	87	74	31	39%
21	77	53	26	40%
22	62	62	24	39%
23	94	76	33	39%
24	85	76	30	37%
25	96	83	41	46%
26	86	76	37	46%
27	80	58	26	38%
28	87	86	33	38%
29	81	70	35	46%
30	79	62	28	40%
31	96	88	42	46%
32	103	92	52	53%
33	46	43	20	45%
34	103	92	52	53%
35	86	87	40	46%
36	83	71	28	36%
37	97	91	37	39%
38	86	71	21	27%
39	87	82	39	46%
40	82	53	28	41%
41	80	88	29	35%
42	90	79	32	38%
43	73	63	34	50%
44	92	65	24	31%
45	85	73	39	49%

Tabulka 15: Úsek 4 čárky

Respondent	Počet čárek	Počet čárek	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	99	84	71	78%
2	76	51	45	71%
3	94	78	63	73%
4	87	74	57	71%
5	94	76	66	78%
6	103	80	67	73%
7	82	79	57	71%
8	101	73	67	77%
9	103	93	71	72%
10	93	73	64	77%
11	94	73	62	74%
12	111	75	68	73%
13	97	72	56	66%
14	106	74	56	62%
15	86	56	44	62%
16	99	76	65	74%
17	93	85	63	71%
18	105	72	68	77%
19	89	69	63	80%
20	91	78	63	75%
21	101	64	59	72%
22	89	76	59	72%
23	106	82	69	73%
24	85	96	60	66%
25	112	86	85	86%
26	97	69	62	75%
27	101	69	53	62%
28	109	89	76	77%
29	92	73	73	88%
30	92	66	59	75%
31	101	83	78	85%
32	108	92	89	89%
33	61	48	42	77%
34	113	92	92	90%
35	86	85	65	76%
36	105	71	66	75%
37	102	81	68	74%
38	98	65	36	44%
39	98	76	74	85%
40	89	61	55	73%
41	95	82	59	67%
42	98	78	61	69%
43	81	65	65	89%
44	123	81	68	67%
45	100	69	64	76%

Tabulka 16: Úsek 5 čárky

Respondent	Počet čárek	Počet čárek	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	107	65	47	55%
2	78	56	46	69%
3	98	69	57	68%
4	89	66	51	66%
5	102	45	41	56%
6	109	68	62	70%
7	96	77	61	71%
8	99	76	71	81%
9	106	87	63	65%
10	110	73	67	73%
11	95	65	51	64%
12	110	79	69	73%
13	96	75	60	70%
14	100	70	52	61%
15	81	54	39	58%
16	109	66	60	69%
17	97	81	64	72%
18	111	73	63	68%
19	99	74	65	75%
20	97	71	60	71%
21	100	59	55	69%
22	101	71	55	64%
23	114	79	66	68%
24	106	83	68	72%
25	102	75	71	80%
26	105	77	68	75%
27	79	73	50	66%
28	110	92	64	63%
29	93	70	68	83%
30	102	74	65	74%
31	103	77	75	83%
32	110	65	52	59%
33	68	53	45	74%
34	119	84	84	83%
35	11	80	11	24%
36	104	70	58	67%
37	112	83	73	75%
38	98	58	26	33%
39	97	73	70	82%
40	92	64	61	78%
41	103	84	70	75%
42	101	85	65	70%
43	87	63	62	83%
44	104	81	56	61%
45	106	78	67	73%

8.1.4 Otazníky

Otazníky jsou doménou tázacích vět, v jiných typech vět se prakticky nevyskytují. Vzhledem k tomu, že námi zkoumaný text je přepisem rozhlasových rozhovorů, očekávala se časté používání interogativních vět v řeči redaktora. Hrubým odhadem by tedy otazníky měly tvořit přibližně třetinu z celkového počtu znamének v textu. Jak se ovšem přesvědčíme v tabulkách na dalších stranách, otazníků bylo doplněno zanedbatelné množství.

Hledejme tedy příčinu tohoto jevu. Pokud pomineme variantu, že respondenti doplnili interpunkční znaménka chybně, je možné, že v českém jazyce se opravdu objevuje zajímavý jev, a to, že k zjišťování informací už se neužívá výhradně tázacích vět, ale že jejich místo zastoupily věty oznamovací s interogativní funkcí. Konkrétní příklad nalezneme v Úseku 3, kde sice nemá otázka formu tázací věty, ale věta oznamovací vybízí k reakci, odpovědi:

„Bezesporu už tato strana svým vlivem na obsazení právě správních rad nakonec mohla odvolávat šéfy zdravotních pojišťoven, to asi nepopřete.“

Další podobný příklad nacházíme v Úseku 1. Věta má sice formu oznamovací věty, na první pohled bychom ji považovali za konstatování. V textu ovšem jasně stojí na pozici otázky a je na ní také odpovídáno.

„Tam si myslíte, že jisté to Poláci nemají, že tam rakety nebo antirakety mít budou.“

„Tak jisté jsou věci teprve, když se stanou, a v tomto smyslu mi nepřísluší žádný komentář.“

Tabulka 17: Úsek 1 otazníky

Respondent	Počet teček	Počet teček	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	9	11	8	80%
2	9	7	5	63%
3	10	5	4	53%
4	11	12	11	96%
5	12	10	10	91%
6	12	9	9	86%
7	10	9	8	84%
8	10	9	8	84%
9	11	11	10	91%
10	11	9	8	80%
11	11	11	10	91%
12	8	8	8	100%
13	11	10	8	76%
14	9	9	6	67%
15	5	7	4	67%
16	12	11	10	87%
17	10	9	6	63%
18	6	9	5	67%
19	12	12	11	92%
20	10	8	8	89%
21	8	5	3	46%
22	11	12	9	78%
23	5	10	5	67%
24	4	2	0	0%
25	9	8	8	94%
26	8	9	6	71%
27	8	12	7	70%
28	11	8	7	74%
29	8	8	8	100%
30	11	9	8	80%
31	10	10	10	100%
32	6	4	4	80%
33	10	10	9	90%
34	6	7	6	92%
35	9	12	8	76%
36	8	9	7	82%
37	10	9	9	95%
38	12	9	9	86%
39	11	11	11	100%
40	7	6	5	77%
41	10	11	9	86%
42	10	8	7	78%
43	11	11	11	100%
44	6	9	5	67%
45	11	12	9	78%

Tabulka 18: Úsek 2 otazníky

Respondent	Počet otazníků	Počet otazníků	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	3	4	1	29%
2	4	2	1	33%
3	5	3	1	25%
4	3	3	1	33%
5	3	3	1	33%
6	6	4	2	40%
7	5	3	2	50%
8	4	4	2	50%
9	4	4	2	50%
10	2	4	1	33%
11	6	5	1	18%
12	4	3	1	29%
13	3	4	1	29%
14	4	3	0	0%
15	2	4	1	33%
16	4	4	2	50%
17	3	3	1	33%
18	2	3	1	40%
19	4	5	2	44%
20	4	2	1	33%
21	3	4	2	57%
22	5	4	2	44%
23	3	4	1	29%
24	2	2	1	50%
25	3	3	1	33%
26	4	3	1	29%
27	2	3	1	40%
28	4	5	2	44%
29	3	3	2	67%
30	4	4	1	25%
31	4	4	1	25%
32	3	3	1	33%
33	5	5	1	20%
34	3	3	1	33%
35	3	3	2	67%
36	2	4	1	33%
37	5	4	2	44%
38	5	1	0	0%
39	5	5	2	40%
40	2	2	1	50%
41	5	4	1	22%
42	4	3	1	29%
43	4	4	2	50%
44	4	2	1	33%
45	5	6	2	36%

Tabulka 19: Úsek 3 otazníky

Respondent	Počet otazníků	Počet otazníků	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	6	4	0	0%
2	5	4	0	0%
3	4	5	0	0%
4	5	5	0	0%
5	4	4	0	0%
6	6	3	0	0%
7	6	4	0	0%
8	7	6	0	0%
9	6	5	0	0%
10	6	5	0	0%
11	6	6	0	0%
12	5	4	0	0%
13	5	5	0	0%
14	6	7	0	0%
15	4	6	1	20%
16	4	5	0	0%
17	6	5	0	0%
18	4	4	0	0%
19	5	4	0	0%
20	5	3	0	0%
21	5	5	0	0%
22	9	6	0	0%
23	6	4	0	0%
24	5	4	0	0%
25	4	4	0	0%
26	6	4	0	0%
27	4	4	0	0%
28	4	4	0	0%
29	5	4	0	0%
30	6	5	0	0%
31	6	5	0	0%
32	6	5	0	0%
33	8	6	0	0%
34	6	5	0	0%
35	5	5	0	0%
36	5	4	0	0%
37	6	5	0	0%
38	6	3	0	0%
39	5	4	0	0%
40	5	2	0	0%
41	6	5	0	0%
42	5	5	0	0%
43	5	4	0	0%
44	5	3	0	0%
45	6	6	0	0%

Tabulka 20: Úsek 4 otazníky

Respondent	Počet otazníků	Počet otazníků	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	6	8	4	57%
2	10	4	1	14%
3	8	6	3	43%
4	7	7	1	14%
5	10	5	5	67%
6	6	6	4	67%
7	8	4	4	67%
8	9	8	6	71%
9	10	7	5	59%
10	10	6	6	75%
11	10	7	4	47%
12	8	4	3	50%
13	9	5	4	57%
14	9	10	4	42%
15	3	5	1	25%
16	10	6	4	50%
17	8	8	4	50%
18	6	5	2	36%
19	9	7	5	63%
20	11	5	4	50%
21	5	4	3	67%
22	7	8	3	40%
23	11	8	6	63%
24	3	2	2	80%
25	11	5	5	63%
26	8	7	4	53%
27	10	7	6	71%
28	9	6	5	67%
29	9	5	5	71%
30	8	7	5	67%
31	10	8	7	78%
32	7	5	5	83%
33	9	9	6	67%
34	8	5	5	77%
35	6	5	4	73%
36	5	5	4	80%
37	10	7	5	59%
38	9	5	3	43%
39	11	8	7	74%
40	10	2	1	17%
41	10	8	4	44%
42	10	8	5	56%
43	8	5	5	77%
44	9	6	5	67%
45	12	10	8	73%

Tabulka 21: Úsek 5 otazníky

Respondent	Počet otazníků	Počet otazníků	Počet shod	% shody
	Text	Zvuk	Text + Zvuk	
1	4	3	2	57%
2	5	2	2	57%
3	7	2	2	44%
4	6	3	2	44%
5	4	3	2	57%
6	8	3	3	55%
7	5	2	2	57%
8	6	3	3	67%
9	5	2	2	57%
10	6	3	3	67%
11	6	3	3	67%
12	4	2	2	67%
13	4	4	3	75%
14	5	4	3	67%
15	3	4	2	57%
16	5	3	3	75%
17	4	3	3	86%
18	3	2	1	40%
19	4	3	3	86%
20	5	3	3	75%
21	5	2	1	29%
22	6	5	3	55%
23	5	3	3	75%
24	3	1	1	50%
25	6	2	2	50%
26	5	3	3	75%
27	5	3	3	75%
28	6	3	3	67%
29	6	4	4	80%
30	4	4	3	75%
31	3	3	2	67%
32	6	3	3	67%
33	5	3	3	75%
34	5	3	3	75%
35	2	3	2	80%
36	3	3	2	67%
37	6	3	3	67%
38	7	3	3	60%
39	5	3	3	75%
40	6	4	4	80%
41	6	3	3	67%
42	5	2	2	57%
43	5	3	3	75%
44	6	3	3	67%
45	5	4	3	67%

8.1.5 Vykřičníky

Výskyt vykřičníků v doplněném textu byl ojedinělý. Mimo fakt, že vykřičníky jsou nejméně užívaným interpunkčním znaménkem z námi zkoumaných čtyř variant, důležitou roli v otázce skromného užití vykřičníků sehrál také charakter posuzovaného textu, který svým laděním nedával žádný nebo naprosto minimální prostor pro imperativní projevy.

Tabulka 22 obsahuje data ze všech úseků, řazena následujícím způsobem. První sloupec informuje o úseku (*Úsek*), k němuž se data vztahují, ve druhém sloupci jsou zobrazena čísla respondentů (*Respondent*), v jejichž textové nebo zvukové podobě se objevily vykřičníky. Čísla ostatních respondentů chybí, protože ani v jednou typu textu nedoplnili žádný vykřičník. Další sloupce obsahují počty doplněných vykřičníků (*Počet vykřičníků*) v textové podobě (*Text*) a zvukové podobě (*Zvuk*), následuje počet shod v umístění vykřičníků na jednom místě v textu (*Počet shod Text + Zvuk*) a procentuálně vyjádřená shoda (*% shody*).

Celkem se v opravených textech objevilo 53 otazníků v Textu a 31 otazníků ve Zvuku. Z tohoto množství se otazník v Textu a Zvuku objevil na stejném místě 18 krát, což činí shodu 34 %.

Jak můžeme z tabulky vyčíst, nejméně vykřičníků bylo respondenty doplněno do Úseku 2. Pouze dva respondenti z celkového počtu 45 respondentů využili možnost použít vykřičník (respondenti s čísly 21 a 45), a to jen 3 krát.

Naopak nejvíce vykřičníků najdeme v Úseku 4, kde toto interpunkční znaménko zvolilo celkem 12 respondentů, a to 27 krát v Textu a 15 krát ve Zvuku. Shoda byla v tomto případě zaznamenána v 11 případech, procentuálně se tedy respondenti shodli z 51 %.

Tabulka 22: Úsek 1-5 vykřičníky

Úsek	Respondent	Počet vykřičníků		Počet shod	% shody
		Text	Zvuk	Text + Zvuk	
1	6	2	2	1	50%
	17	2	0	0	0%
	21	0	1	0	0%
	31	3	1	1	50%
	34	1	1	1	100%
	41	0	1	0	0%
	45	2	1	0	0%
2	21	1	0	0	0%
	45	1	1	1	100%
3	6	1	0	0	0%
	9	1	0	0	0%
	21	1	0	0	0%
	38	1	0	0	0%
	45	0	1	0	0%
4	6	5	4	3	67%
	10	1	0	0	0%
	21	3	1	0	0%
	31	1	1	1	100%
	32	1	1	1	100%
	33	2	1	1	67%
	34	3	2	2	80%
	35	2	2	2	100%
	38	3	0	0	0%
	39	1	1	1	100%
	40	0	2	0	0%
	45	5	0	0	0%
5	6	2	2	1	50%
	17	2	0	0	0%
	21	0	1	0	0%
	31	3	1	1	50%
	34	1	1	1	100%
	41	0	1	0	0%
	45	2	1	0	0%
Součty		53	31	18	34%

8.2 Porovnání s automatikou

Tento oddíl se vypořádává s otázkou automatického doplňování interpunkce speciálním programem a aplikovatelností této metody na počítačem rozpoznáný text. Počítači byl ke korektuře dodán stejný text, jímž se zabývali respondenti. Předmětem zkoumání je shoda v počtu znaků, která by měla ukázat, jestli se automatika shoduje spíše s doplněním čárek do textové, nebo do zvukové podoby.

Předpokládáme, že vzhledem k tomu, že automatika pracovala pouze se základními gramatickými dovednostmi a při umísťování čárek na nejednoznačná místa hrála hlavní roli pravděpodobnost vycházející z dat z korpusu, je pravděpodobné, že se bude shodovat spíše s textovou verzí, protože při doplňování do zvukové podoby měli respondenti k dispozici zatím technikou nepřenositelné informace – mimo jazykového citu a možnosti výběru logicky zdůvodnitelnější varianty to byla především intonace. Při výběru interpunkce pro Text se ale museli řídit obdobnými zákonitostmi jako automatika.

8.2.1 Automatika obecně

Tabulka 23: Automatika

Automatika				
Úsek	Tečky	Čárky	Vykřičníky	Otazníky
Úsek 1	35	98	0	9
Úsek 2	29	128	0	5
Úsek 3	26	115	1	9
Úsek 4	43	130	0	10
Úsek 5	45	141	0	8
Součet	178	612	1	41

V Tabulce 23 se dočítáme, že automatika doplnila do Textu celkem 178 teček, oproti tomu respondenti (Tabulka 24) doplnili 159 teček do Textu a 129 teček do Zvuku, menší rozdíl je tedy mezi dvojicí automatika – respondenti Text.

Stejného výsledku se dobereme v oblasti čárek: automatika 612 čárek, respondenti Text 423 čárek, respondenti Zvuk 364 čárek. Potvrdila se tedy domněnka,

že informace, které respondenti získali z intonace, je pro doplnění chybějící interpunkce klíčová a program si nevystačí pouze se znalostí gramatiky.

Vzhledem k velkému rozptýlu dat je k vynesení jasného verdiktu nutné provést ještě další zkoumání.

Tabulka 24: Respondenti

Respondenti								
Úsek	Tečky		Čárky		Vykřičníky		Otazníky	
	Text	Zvuk	Text	Zvuk	Text	Zvuk	Text	Zvuk
Úsek 1	31	29	62	61	0	0	9	9
Úsek 2	29	29	84	84	0	0	4	4
Úsek 3	23	17	83	72	0	0	5	5
Úsek 4	40	30	96	75	1	0	8	6
Úsek 5	36	24	98	72	0	0	5	3
Součet	159	129	423	364	1	0	31	27

8.2.2 Automatika konkrétně

Tabulky 25-30 komentují shody mezi automaticky řešenou interpunkcí a texty dodanými respondenty.

V prvním sloupci se nachází číslo respondenta (*Respondent*), ve druhém (*Shoda / rozdíl*) získáváme informaci, zda na daném řádku řešíme shodu, nebo rozdíl mezi automatikou a respondenty. Ve sloupci nazvaném *Text* je uveden počet znamének, v nichž se respondent shodl s automatikou doplněnými znaménky, další sloupec (*% shody Text – automatika*) nám tento počet vyjadřuje procentuálně. Sloupec *Zvuk* obsahuje informace o počtu shod zvukové stopy respondenta s automaticky doplněnou interpunkcí a poslední sloupec uzavírá opět procentuální hodnotou shody (*% shody Zvuk – automatika*).

Například respondent číslo 6 se s automatikou, která do Úseku 1 doplnila celkem 142 znamének, shodl v 86 případech (61 %) v Textu a 80 případech (56 %) u Zvuku, neshodl se v 15 případech (11 %) u Textu a 15 případech (11 %) u Zvuku.

Tabulka 25: Úsek 1: Porovnání automatiky a respondentů

Respondent	Shoda / rozdíl	Text	% shody Text - automatika	Zvuk	% shody Zvuk - automatika
1	shoda	128	90%	118	83%
	rozdíl	8	6%	17	12%
2	shoda	72	51%	70	49%
	rozdíl	12	8%	17	12%
3	shoda	94	66%	70	49%
	rozdíl	15	11%	12	8%
4	shoda	82	58%	76	54%
	rozdíl	16	11%	13	9%
5	shoda	86	61%	87	61%
	rozdíl	15	11%	14	10%
6	shoda	86	61%	80	56%
	rozdíl	15	11%	15	11%
7	shoda	80	56%	88	62%
	rozdíl	10	7%	14	10%
8	shoda	92	65%	91	64%
	rozdíl	15	11%	11	8%
9	shoda	86	61%	93	65%
	rozdíl	11	8%	13	9%
10	shoda	99	70%	90	63%
	rozdíl	12	8%	16	11%
11	shoda	81	57%	79	56%
	rozdíl	19	13%	16	11%
12	shoda	89	63%	89	63%
	rozdíl	13	9%	15	11%
13	shoda	72	51%	75	53%
	rozdíl	23	16%	15	11%
14	shoda	93	65%	85	60%
	rozdíl	9	6%	19	13%
15	shoda	77	54%	86	61%
	rozdíl	24	17%	9	6%
16	shoda	89	63%	83	58%
	rozdíl	18	13%	21	15%
17	shoda	89	63%	79	56%
	rozdíl	18	13%	22	15%
18	shoda	81	57%	86	61%
	rozdíl	20	14%	13	9%
19	shoda	91	64%	86	61%
	rozdíl	11	8%	13	9%
20	shoda	89	63%	87	61%
	rozdíl	15	11%	15	11%
21	shoda	72	51%	68	48%
	rozdíl	17	12%	25	18%
22	shoda	94	66%	78	55%
	rozdíl	14	10%	18	13%

23	shoda	90	63%	87	61%
	rozdíl	20	14%	17	12%
24	shoda	77	54%	67	47%
	rozdíl	18	13%	31	22%
25	shoda	92	65%	91	64%
	rozdíl	13	9%	15	11%
26	shoda	100	70%	97	68%
	rozdíl	14	10%	12	8%
27	shoda	74	52%	77	54%
	rozdíl	20	14%	17	12%
28	shoda	91	64%	68	48%
	rozdíl	15	11%	23	16%
29	shoda	92	65%	92	65%
	rozdíl	14	10%	13	9%
30	shoda	91	64%	85	60%
	rozdíl	11	8%	13	9%
31	shoda	83	58%	88	62%
	rozdíl	14	10%	15	11%
32	shoda	93	65%	89	63%
	rozdíl	19	13%	18	13%
33	shoda	76	54%	70	49%
	rozdíl	14	10%	14	10%
34	shoda	92	65%	94	66%
	rozdíl	18	13%	20	14%
35	shoda	95	67%	79	56%
	rozdíl	11	8%	11	8%
36	shoda	81	57%	79	56%
	rozdíl	22	15%	20	14%
37	shoda	99	70%	94	66%
	rozdíl	13	9%	15	11%
38	shoda	91	64%	56	39%
	rozdíl	14	10%	27	19%
39	shoda	85	60%	93	65%
	rozdíl	15	11%	14	10%
40	shoda	94	66%	76	54%
	rozdíl	15	11%	16	11%
41	shoda	88	62%	84	59%
	rozdíl	11	8%	21	15%
42	shoda	96	68%	86	61%
	rozdíl	16	11%	19	13%
43	shoda	92	65%	92	65%
	rozdíl	13	9%	12	8%
44	shoda	83	58%	74	52%
	rozdíl	12	8%	17	12%
45	shoda	97	68%	86	61%
	rozdíl	22	15%	23	16%

Tabulka 26: Úsek 2: Porovnání automatiky a respondentů

Respondent	Shoda / rozdíl	Text	% shody Text - automatika	Zvuk	% shody Zvuk - automatika
1	shoda	96	59%	73	45%
	rozdíl	28	17%	24	15%
2	shoda	73	45%	71	44%
	rozdíl	23	14%	31	19%
3	shoda	94	58%	94	58%
	rozdíl	33	20%	31	19%
4	shoda	83	51%	77	48%
	rozdíl	32	20%	30	19%
5	shoda	91	56%	92	57%
	rozdíl	22	14%	25	15%
6	shoda	79	49%	84	52%
	rozdíl	29	18%	30	19%
7	shoda	68	42%	92	57%
	rozdíl	30	19%	28	17%
8	shoda	91	56%	89	55%
	rozdíl	29	18%	36	22%
9	shoda	89	55%	91	56%
	rozdíl	26	16%	39	24%
10	shoda	89	55%	85	52%
	rozdíl	25	15%	33	20%
11	shoda	86	53%	75	46%
	rozdíl	24	15%	35	22%
12	shoda	94	58%	91	56%
	rozdíl	22	14%	26	16%
13	shoda	79	49%	68	42%
	rozdíl	27	17%	23	14%
14	shoda	85	52%	89	55%
	rozdíl	32	20%	32	20%
15	shoda	75	46%	75	46%
	rozdíl	42	26%	40	25%
16	shoda	84	52%	81	50%
	rozdíl	27	17%	28	17%
17	shoda	98	60%	101	62%
	rozdíl	41	25%	35	22%
18	shoda	87	54%	88	54%
	rozdíl	31	19%	30	19%
19	shoda	89	55%	81	50%
	rozdíl	29	18%	31	19%
20	shoda	86	53%	87	54%
	rozdíl	34	21%	27	17%
21	shoda	81	50%	71	44%
	rozdíl	28	17%	25	15%
22	shoda	84	52%	82	51%
	rozdíl	29	18%	39	24%

23	shoda	96	59%	94	58%
	rozdíl	33	20%	33	20%
24	shoda	82	51%	77	48%
	rozdíl	27	17%	34	21%
25	shoda	92	57%	91	56%
	rozdíl	26	16%	27	17%
26	shoda	95	59%	88	54%
	rozdíl	29	18%	32	20%
27	shoda	87	54%	77	48%
	rozdíl	26	16%	21	13%
28	shoda	87	54%	78	48%
	rozdíl	30	19%	41	25%
29	shoda	100	62%	100	62%
	rozdíl	32	20%	32	20%
30	shoda	84	52%	88	54%
	rozdíl	25	15%	30	19%
31	shoda	84	52%	95	59%
	rozdíl	32	20%	31	19%
32	shoda	87	54%	87	54%
	rozdíl	34	21%	34	21%
33	shoda	71	44%	69	43%
	rozdíl	31	19%	30	19%
34	shoda	87	54%	87	54%
	rozdíl	34	21%	34	21%
35	shoda	95	59%	63	39%
	rozdíl	25	15%	33	20%
36	shoda	85	52%	91	56%
	rozdíl	28	17%	30	19%
37	shoda	96	59%	92	57%
	rozdíl	35	22%	32	20%
38	shoda	89	55%	64	40%
	rozdíl	27	17%	36	22%
39	shoda	91	56%	93	57%
	rozdíl	23	14%	21	13%
40	shoda	85	52%	95	59%
	rozdíl	33	20%	29	18%
41	shoda	90	56%	97	60%
	rozdíl	39	24%	32	20%
42	shoda	90	56%	98	60%
	rozdíl	28	17%	30	19%
43	shoda	89	55%	89	55%
	rozdíl	31	19%	31	19%
44	shoda	93	57%	82	51%
	rozdíl	24	15%	30	19%
45	shoda	90	56%	89	55%
	rozdíl	42	26%	42	26%

Tabulka 27: Úsek 3: Porovnání automatiky a respondentů

Respondent	Shoda / rozdíl	Text	% shody Text - automatika	Zvuk	% shody Zvuk - automatika
1	shoda	97	64%	66	44%
	rozdíl	19	13%	19	13%
2	shoda	79	52%	63	42%
	rozdíl	15	10%	19	13%
3	shoda	98	65%	84	56%
	rozdíl	24	16%	24	16%
4	shoda	70	46%	77	51%
	rozdíl	17	11%	13	9%
5	shoda	80	53%	77	51%
	rozdíl	22	15%	16	11%
6	shoda	87	58%	64	42%
	rozdíl	23	15%	16	11%
7	shoda	84	56%	72	48%
	rozdíl	15	10%	17	11%
8	shoda	93	62%	77	51%
	rozdíl	20	13%	19	13%
9	shoda	102	68%	89	59%
	rozdíl	20	13%	27	18%
10	shoda	98	65%	81	54%
	rozdíl	18	12%	16	11%
11	shoda	78	52%	66	44%
	rozdíl	19	13%	14	9%
12	shoda	93	62%	76	50%
	rozdíl	21	14%	13	9%
13	shoda	82	54%	66	44%
	rozdíl	13	9%	9	6%
14	shoda	95	63%	78	52%
	rozdíl	20	13%	10	7%
15	shoda	83	55%	70	46%
	rozdíl	22	15%	16	11%
16	shoda	92	61%	72	48%
	rozdíl	16	11%	11	7%
17	shoda	111	74%	89	59%
	rozdíl	18	12%	20	13%
18	shoda	92	61%	87	58%
	rozdíl	16	11%	19	13%
19	shoda	98	65%	81	54%
	rozdíl	16	11%	16	11%
20	shoda	97	64%	75	50%
	rozdíl	19	13%	16	11%
21	shoda	76	50%	57	38%
	rozdíl	21	14%	12	8%
22	shoda	81	54%	68	45%

	rozdíl	10	7%	15	10%
23	shoda	99	66%	82	54%
	rozdíl	22	15%	16	11%
24	shoda	82	54%	70	46%
	rozdíl	23	15%	18	12%
25	shoda	99	66%	85	56%
	rozdíl	22	15%	17	11%
26	shoda	105	70%	87	58%
	rozdíl	17	11%	16	11%
27	shoda	85	56%	61	40%
	rozdíl	19	13%	13	9%
28	shoda	95	63%	78	52%
	rozdíl	20	13%	21	14%
29	shoda	97	64%	80	53%
	rozdíl	17	11%	16	11%
30	shoda	99	66%	79	52%
	rozdíl	16	11%	15	10%
31	shoda	106	70%	93	62%
	rozdíl	16	11%	15	10%
32	shoda	108	72%	91	60%
	rozdíl	18	12%	17	11%
33	shoda	76	50%	61	40%
	rozdíl	12	8%	12	8%
34	shoda	108	72%	91	60%
	rozdíl	18	12%	17	11%
35	shoda	97	64%	88	58%
	rozdíl	20	13%	23	15%
36	shoda	85	56%	71	47%
	rozdíl	18	12%	18	12%
37	shoda	108	72%	89	59%
	rozdíl	20	13%	22	15%
38	shoda	104	69%	67	44%
	rozdíl	14	9%	32	21%
39	shoda	103	68%	83	55%
	rozdíl	17	11%	19	13%
40	shoda	96	64%	63	42%
	rozdíl	19	13%	16	11%
41	shoda	106	70%	89	59%
	rozdíl	18	12%	17	11%
42	shoda	101	67%	82	54%
	rozdíl	16	11%	18	12%
43	shoda	88	58%	71	47%
	rozdíl	13	9%	13	9%
44	shoda	91	60%	64	42%
	rozdíl	20	13%	17	11%
45	shoda	103	68%	88	58%
	rozdíl	15	10%	17	11%

Tabulka 28: Úsek 4: Porovnání automatiky a respondentů

Respondent	Shoda / rozdíl	Text	% shody Text - automatika	Zvuk	% shody Zvuk - automatika
1	shoda	109	60%	91	50%
	rozdíl	28	15%	27	15%
2	shoda	108	59%	76	42%
	rozdíl	18	10%	22	12%
3	shoda	129	70%	100	55%
	rozdíl	23	13%	30	16%
4	shoda	104	57%	87	48%
	rozdíl	19	10%	27	15%
5	shoda	120	66%	89	49%
	rozdíl	25	14%	22	12%
6	shoda	115	63%	83	45%
	rozdíl	34	19%	30	16%
7	shoda	113	62%	84	46%
	rozdíl	17	9%	26	14%
8	shoda	128	70%	85	46%
	rozdíl	23	13%	24	13%
9	shoda	126	69%	109	60%
	rozdíl	28	15%	28	15%
10	shoda	120	66%	90	49%
	rozdíl	26	14%	19	10%
11	shoda	107	58%	89	49%
	rozdíl	32	17%	24	13%
12	shoda	121	66%	84	46%
	rozdíl	29	16%	22	12%
13	shoda	102	56%	77	42%
	rozdíl	28	15%	21	11%
14	shoda	128	70%	84	46%
	rozdíl	28	15%	32	17%
15	shoda	121	66%	79	43%
	rozdíl	26	14%	26	14%
16	shoda	123	67%	84	46%
	rozdíl	26	14%	24	13%
17	shoda	122	67%	102	56%
	rozdíl	31	17%	25	14%
18	shoda	119	65%	89	49%
	rozdíl	28	15%	18	10%
19	shoda	123	67%	81	44%
	rozdíl	20	11%	28	15%
20	shoda	122	67%	78	43%
	rozdíl	20	11%	27	15%
21	shoda	95	52%	69	38%
	rozdíl	37	20%	24	13%
22	shoda	111	61%	83	45%
	rozdíl	28	15%	29	16%

23	shoda	127	69%	95	52%
	rozdíl	28	15%	27	15%
24	shoda	91	50%	81	44%
	rozdíl	25	14%	29	16%
25	shoda	133	73%	95	52%
	rozdíl	25	14%	24	13%
26	shoda	126	69%	91	50%
	rozdíl	20	11%	22	12%
27	shoda	111	61%	73	40%
	rozdíl	34	19%	21	11%
28	shoda	123	67%	86	47%
	rozdíl	29	16%	27	15%
29	shoda	122	67%	90	49%
	rozdíl	23	13%	21	11%
30	shoda	120	66%	92	50%
	rozdíl	28	15%	19	10%
31	shoda	126	69%	95	52%
	rozdíl	27	15%	22	12%
32	shoda	118	64%	95	52%
	rozdíl	28	15%	26	14%
33	shoda	100	55%	76	42%
	rozdíl	31	17%	19	10%
34	shoda	125	68%	95	52%
	rozdíl	31	17%	26	14%
35	shoda	118	64%	102	56%
	rozdíl	21	11%	23	13%
36	shoda	110	60%	75	41%
	rozdíl	29	16%	21	11%
37	shoda	121	66%	93	51%
	rozdíl	37	20%	27	15%
38	shoda	119	65%	60	33%
	rozdíl	30	16%	37	20%
39	shoda	123	67%	90	49%
	rozdíl	26	14%	24	13%
40	shoda	125	68%	77	42%
	rozdíl	22	12%	23	13%
41	shoda	126	69%	86	47%
	rozdíl	32	17%	30	16%
42	shoda	125	68%	88	48%
	rozdíl	27	15%	29	16%
43	shoda	115	63%	85	46%
	rozdíl	23	13%	19	10%
44	shoda	135	74%	88	48%
	rozdíl	28	15%	25	14%
45	shoda	124	68%	92	50%
	rozdíl	35	19%	23	13%

Tabulka 29: Úsek 5: Porovnání automatiky a respondentů

Respondent	Shoda / rozdíl	Text	% shody Text - automatika	Zvuk	% shody Zvuk - automatika
1	shoda	110	57%	73	38%
	rozdíl	31	16%	24	12%
2	shoda	103	53%	72	37%
	rozdíl	22	11%	22	11%
3	shoda	118	61%	89	46%
	rozdíl	29	15%	23	12%
4	shoda	98	51%	83	43%
	rozdíl	24	12%	14	7%
5	shoda	116	60%	53	27%
	rozdíl	24	12%	13	7%
6	shoda	122	63%	80	41%
	rozdíl	29	15%	19	10%
7	shoda	116	60%	83	43%
	rozdíl	16	8%	17	9%
8	shoda	125	64%	89	46%
	rozdíl	22	11%	21	11%
9	shoda	117	60%	94	48%
	rozdíl	26	13%	21	11%
10	shoda	132	68%	82	42%
	rozdíl	17	9%	16	8%
11	shoda	111	57%	82	42%
	rozdíl	28	14%	20	10%
12	shoda	118	61%	81	42%
	rozdíl	25	13%	19	10%
13	shoda	99	51%	74	38%
	rozdíl	24	12%	20	10%
14	shoda	119	61%	73	38%
	rozdíl	23	12%	21	11%
15	shoda	97	50%	73	38%
	rozdíl	37	19%	25	13%
16	shoda	125	64%	72	37%
	rozdíl	25	13%	20	10%
17	shoda	124	64%	88	45%
	rozdíl	36	19%	26	13%
18	shoda	110	57%	85	44%
	rozdíl	32	16%	16	8%
19	shoda	117	60%	87	45%
	rozdíl	22	11%	14	7%
20	shoda	120	62%	77	40%
	rozdíl	20	10%	13	7%
21	shoda	103	53%	62	32%
	rozdíl	26	13%	17	9%
22	shoda	118	61%	79	41%
	rozdíl	30	15%	17	9%

23	shoda	121	62%	86	44%
	rozdíl	30	15%	21	11%
24	shoda	101	52%	71	37%
	rozdíl	28	14%	24	12%
25	shoda	125	64%	83	43%
	rozdíl	20	10%	19	10%
26	shoda	124	64%	91	47%
	rozdíl	29	15%	19	10%
27	shoda	88	45%	72	37%
	rozdíl	22	11%	20	10%
28	shoda	123	63%	81	42%
	rozdíl	25	13%	26	13%
29	shoda	118	61%	86	44%
	rozdíl	23	12%	18	9%
30	shoda	115	59%	84	43%
	rozdíl	26	13%	25	13%
31	shoda	115	59%	83	43%
	rozdíl	27	14%	18	9%
32	shoda	123	63%	82	42%
	rozdíl	25	13%	20	10%
33	shoda	103	53%	69	36%
	rozdíl	23	12%	19	10%
34	shoda	125	64%	83	43%
	rozdíl	28	14%	25	13%
35	shoda	15	8%	77	40%
	rozdíl	4	2%	21	11%
36	shoda	103	53%	75	39%
	rozdíl	28	14%	16	8%
37	shoda	130	67%	89	46%
	rozdíl	33	17%	24	12%
38	shoda	124	64%	48	25%
	rozdíl	28	14%	30	15%
39	shoda	125	64%	85	44%
	rozdíl	19	10%	17	9%
40	shoda	120	62%	79	41%
	rozdíl	20	10%	18	9%
41	shoda	126	65%	86	44%
	rozdíl	34	18%	22	11%
42	shoda	120	62%	85	44%
	rozdíl	28	14%	24	12%
43	shoda	115	59%	80	41%
	rozdíl	18	9%	14	7%
44	shoda	106	55%	73	38%
	rozdíl	28	14%	26	13%
45	shoda	126	65%	91	47%
	rozdíl	32	16%	19	10%

9 Závěr

Srovnáními, prezentovanými na předchozích stránkách, se jasně ukázalo, že intonace přidává do rozhodování o umístění interpunkce další informace. Na základě těchto informací se pak respondenti, potažmo všichni uživatelé jazyka, rozhodují, kam doplní, nebo nedoplní interpunkční znaménka, pokud mají zaznamenat mluvený projev.

Při porovnávání dat od respondentů a textu opraveného automatickou bylo zřejmé, že intonace mluvcích je nedílnou součástí informace o interpunkčních znaménkách a že jim přesně odpovídá.

Bakalářská práce poskytuje velké množství sebraných, rozříděných a prvotně zhodnocených dat. Všechna data je možné využít k dalšímu zkoumání, možnosti nebyly zdaleka vyčerpány.

Zajímavým námětem se zdá být výzkum míst, na kterých se respondenti často neshodovali. Některé z důvodů, proč tomu tak mohlo být, jsou nastíněny v předchozích kapitolách, ovšem výzkum zaměřený na části s velmi nízkou shodou, orientovaný i k bezprostřednímu okolí chybějících nebo přebývajících znamének by mohl ukázat, proč činilo respondentům problém se rozhodnout, kam znaménko umístit a co je nakonec vedlo k tomu, že znaménko umístili, či nikoliv.

Prostor k dalšímu zkoumání nabízí také oblast vykřičníků. Jak bylo zmíněno výše, typ rozhlasové debaty s nekonfliktními mluvcími neumožňuje sběr dat s imperativní funkcí. Při jiné volbě výchozího textu by ovšem mohla být srovnána četnost užívání v jednotlivých typech textů.

10 Seznam literatury

- ADAM, Robert. Matná světla na periférii české interpunkce. *Bohemistika*. 2008, roč. 8, s. 339-353. ISSN 1642-9893.
- ČECHOVÁ, Marie a kol.: *Čeština - řeč a jazyk*. Praha : ISV, 2000. ISBN 978-80-7235-413-9.
- GREPL, Miroslav a kol.: *Příruční mluvnice češtiny*. Praha : NLN, 2008. ISBN 978-80-7106-624-8.
- GREPL, Miroslav; KARLÍK, Petr: *Skladba češtiny*. Olomouc : Votobia, 1998. ISBN 80-85866-47-1.
- GREPL, Miroslav; KARLÍK, Petr: *Skladba spisovné češtiny*. Praha : SPN, 1989. ISBN 14-608-86
- HUBÁČEK, Jaroslav a kol.: *Čeština pro učitele*. Opava : Vademecum, 2002. ISBN 978-80-86041-37-7.
- KARLÍK, Petr; NEKULA, Marek; PLESKALOVÁ, Jana. *Encyklopedický slovník češtiny*. Praha : Nakladatelství Lidové noviny, 2002. ISBN 80-7106-484-3.
- Kolektiv autorů Ústavu českého jazyka Filosofické fakulty Masarykovy univerzity v Brně, edit. KARLÍK, Petr; NEKULA, Marek; RUSÍNOVÁ, Zdenka. *Příruční mluvnice češtiny*. Praha : Nakladatelství Lidové noviny, 2008. ISBN 978-80-7106-624-8.
- NOVOTNÝ, Jiří: *Nástin syntaktického popisu češtiny (na základě valenční teorie)*. Ústí na Labem : Univerzita J.E.Purkyně, 1994. ISBN: 8070440864.
- PALKOVÁ, Zdena: *Fonetika a fonologie češtiny : s obecným úvodem do problematiky oboru*. Praha : Karolinum, 1997. ISBN 80-7066-843-1.
- SEDLÁČEK, Miloslav. K základním otázkám interpunkce v češtině. *Naše řeč*. 1986, roč. 69, č. 3, s. 121-132. ISSN 0027-8203.
- ŠMILAUER, Vladimír: *Novočeská skladba*. Praha : SPN, 1966. ISBN 80-7198-281-4.

11 Přílohy

Nedílnou součástí bakalářské práce je přiložené CD, na kterém se nachází sebraná data od respondentů (textová podoba, zvuková podoba, Profil respondenta), automatický přepis, programy Nanotrans a Transcript Viewer a celá bakalářská práce ve formátu pdf.

11.1 Průvodce

Průvodce

Vaším úkolem bude doplnění interpunkčních znamének (čárek, teček, středníků atd.) do textu, z něhož byla všechna znaménka záměrně odstraněna. Úkol budete zpracovávat ve speciálním programu NANOTRANS.

Vysvětlivky

- text je přepisem mluveného slova, proto se v určitých místech objevují symboly, které upozorňují na přechyby a nedorečenosti, jichž se mluvčí dopustil
 - svislé čáry v textu |něco| → přechyby
 - poznámky v hranatých závorkách [něco] → mimojazykové jevy (kašel, smích)

Průvodce jednotlivými kroky

Velmi podrobný, nemějte strach z množství úkolů! ☺

- 1) Otevřete složku „*NanoTrans*“ a spusťte program dvojklikem na „*NanoTrans*“ s ikonkou pera
- 2) V horní liště klikněte na „*File*“ a vyberte „*Open Transcription*“
- 3) Otevřete adresář „*Doplnění interpunkce*“, rozklikněte složku „*NanoTrans*“, pokračujte na „*Data*“ a vložte do programu text dvojklikem na „*Přepis bez interpunkce.trxs*“
- 4) Zvolte znovu „*File*“, pokračujte na „*Save as*“ a uložte na snadno dohledatelné místo na Vašem disku textový soubor (přípona *.trxs nebo *.html) pod název „*Text*“
- 5) Do textu doplňte interpunkční znaménka

! Průběžně ukládejte stiskem kláves CTRL a S nebo „File“ → „Save“ !

- 6) Po dokončení práce odešlete textový soubor „*Text.trxs*“ na mail hana.plocova@gmail.com

Po uplynutí nějakého času, optimální jsou 3 dny, doplňte čárky znovu, tentokrát ale s pomocí zvukové stopy:

- 1) Otevřete složku „*NanoTrans*“ a spusťte program dvojklikem na „*NanoTrans*“ s ikonkou pera
- 2) V horní liště klikněte na „*File*“ a vyberte „*Open Transcription*“
- 3) Otevřete adresář „*Doplnění interpunkce*“, rozklikněte složku „*NanoTrans*“, pokračujte na „*Data*“ a vložte do programu text dvojklikem na „*Přepis bez interpunkce.trxs*“
- 4) Zvolte znovu „*File*“, pokračujte na „*Save as*“ a uložte na snadno dohledatelné místo na Vašem disku textový soubor (přípona *.trxs nebo *.html) pod název „*Zvuk*“
- 5) Stiskem klávesy TAB spustíte zvukový záznam (dole ve spodní části můžete sledovat průběh přehrávání), opakovaným stiskem přehrávání zastavíte
- 6) Podržením klávesy CTRL a klikem levého tlačítka myši můžete vybrat přesné místo, odkud chcete přehrávání začít
- 7) Po zastavení přehrávání zvuku (stisk TAB) se můžete podržením klávesy ALT a stiskem příslušné šipky pohybovat o 1 vteřinu zpět či vpřed
- 8) S pomocí zvukové stopy do textu doplňte interpunkční znaménka

! Průběžně ukládejte stiskem kláves CTRL a S nebo „File“ → „Save“ !

- 9) Po dokončení práce odešlete textový soubor „*Zvuk.trxs*“ spolu s vyplněným „*Profilem respondenta*“ (naleznete v hlavní složce „*Doplnění interpunkce*“) na mail hana.plocova@gmail.com

Velice a srdečně děkujeme za spolupráci!

V případě jakýchkoli dotazů se obraťte na hana.plocova@gmail.com

11.2 Profil respondenta

Profil respondenta

Prosíme vyplňte, případně zakroužkujte:

Pohlaví

Muž

Žena

Věk

....

Nejvyšší dosažené vzdělání

Základní

Střední bez maturity

Střední s maturitou

Všeobecné

Odborné

Vysokoškolské

Technické

Humanitní